Check for updates

Received: 14 January 2022 | Revised: 3 June 2022 | Accepted: 16 June 2022

DOI: 10.1111/1755-0998.13680

MOLECULAR ECOLOGY RESOURCES WILEY

RESOURCE ARTICLE

Parent-offspring inference in inbred populations

Jan-Niklas Runge[1,2] | Barbara König[2] | Anna K. Lindholm[2] | Andres Bendesky[1]

[1]Department of Ecology, Evolution and Environmental Biology, Zuckerman Mind Brain Behavior Institute, Columbia University, New York, NY, USA

[2]Department of Evolutionary Biology and Environmental Studies, University of Zurich, Zürich, Switzerland

Correspondence
Jan-Niklas Runge and Andres Bendesky, Department of Ecology, Evolution and Environmental Biology, Zuckerman Mind Brain Behavior Institute, Columbia University, New York, NY, USA.
Email: jn.runge@protonmail.com (J.R.); a.bendesky@columbia.edu (A.B.)

Funding information
Klingenstein-Simons Fellowship; National Institutes of Health, Grant/Award Number: R35GM143051; Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung, Grant/Award Number: 31003A_120444, 31003A_160328, 31003A_176114, P1ZHP3_181303 and P2ZHP3_195249; Searle Scholarship; Sloan Foundation Fellowship

Handling Editor: Andrew DeWoody

Abstract

Genealogical relationships are fundamental components of genetic studies. However, it is often challenging to infer correct and complete pedigrees even when genome-wide information is available. For example, inbreeding can obscure genetic differences between individuals, making it difficult to even distinguish first-degree relatives such as parent-offspring from full siblings. Similarly, genotyping errors can interfere with the detection of genetic similarity between parents and their offspring. Inbreeding is common in natural, domesticated, and experimental populations and genotyping of these populations often has more errors than in human data sets, so efficient methods for building pedigrees under these conditions are necessary. Here, we present a new method for parent-offspring inference in inbred pedigrees called specific parent-offspring relationship estimation (SPORE). SPORE is vastly superior to existing pedigree-inference methods at detecting parent-offspring relationships, in particular when inbreeding is high or in the presence of genotyping errors, or both. SPORE therefore fills an important void in the arsenal of pedigree inference tools.

KEYWORDS
genotyping errors, homozygosity, identity by descent, inbreeding, pedigree reconstruction, relatedness

1 | INTRODUCTION

Genealogical relationships among individuals are critical for many genetic analyses. For example, in trait-mapping studies based on linkage (Ott et al., 2015) and transmission disequilibrium analyses (e.g., Spielman & Ewens, 1998) as well as in sibling-based genome-wide association studies (Howe et al., 2021), the transmission of traits and alleles is analysed within pedigrees to control for environmental similarities and to avoid confounders from population structure. Pedigrees can also be fit into statistical models to estimate the heritabilities of traits (Kaplanis et al., 2018; Wilson et al., 2010). Removing close relatives from analyses can also be important to fulfil analytical assumptions in population genetics (Wang, 2004). Pedigrees are also useful to study the rate of de novo mutations and their impact on traits (Koch et al., 2019). Furthermore, pedigrees themselves can be of considerable interest, for example when analysing differences in lifetime reproductive success (Farquharson et al., 2017) or assortative mating (Grant & Grant, 2008). Thus, accurate and complete pedigree relationships are important to maximize power and accuracy of genetic analyses and for studies of population biology.

Deriving complete and correct pedigrees can be challenging. In monitored wild populations, researchers may use a combination of observed data, such as dates of birth and death, coupled with sparse genotypes produced by technologies such as microsatellites, and employ software tools such as COLONY (Wang, 2004) to build

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2022 The Authors. *Molecular Ecology Resources* published by John Wiley & Sons Ltd.

pedigrees. This approach is usually accurate, but has two major drawbacks. First, documenting dates of birth and death (to exclude impossible parent-offspring matches) can be difficult in some species and contexts. Second, it is impractical to genotype large numbers of microsatellites, decreasing the power to infer genealogical relationships in populations with low genetic diversity.

Current genome-wide methods for pedigree reconstruction are often developed and optimized for outbred human populations and thus compatibility with considerable levels of inbreeding is not the main concern (e.g., Huff et al., 2011). However, some human populations and individuals have high levels of inbreeding, with up to 20% of their genome empirically determined to be in a homozygous state (Ceballos et al., 2018; Keller et al., 2011; Yengo et al., 2019; Figure 1a). Considerable inbreeding is also present in populations of other species, including livestock (Alemu et al., 2021; Talebi et al., 2020; and Methods), small populations of wild animals such as wolves (Kardos et al., 2018) and Devils Hole pupfish (Tian et al., 2021), and even in populations with large census sizes such as house mice (Figure 1a). Inbreeding violates the assumptions of many genealogical-relationship inferring software, yet the performance of these tools in the presence of inbreeding has not been thoroughly evaluated. Inbreeding homogenizes genomes within populations, making it difficult to leverage patterns of genetic variant sharing to distinguish different types of relatives. For example, the fractions of their genomes two individuals share identical-by-descent ("IBD," inherited through a recent common ancestor), in only one of the two homologous chromosomes ("IBD1") or in both homologous chromosomes ("IBD2"), differ between parent-offspring and full-sibling relationships. Hence, these metrics are often used to distinguish such relationships (Li et al., 2014; Staples et al., 2014; Stevens et al., 2011). However, when parents are genetically related, the fraction of the genome that is IBD2 with their offspring is higher than in outbred populations and is thus more similar to full-sibling relationships (Figure 1b). Thus, inbreeding could hinder the accuracy of existing pedigree-reconstruction tools.

Typical genotyping error levels in human studies can also have a large impact on pedigree inference (Smith et al., 2021; Wang, 2010) and these errors might be more prevalent in less well-studied species and populations (Bresadola et al., 2020; Browning & Browning, 2013; Money et al., 2015). Genotyping errors can result in parent-offspring relationships appearing as less related than expected, obscuring the distinction between parent-offspring and other types of relationships. Furthermore, these errors can interrupt inferred IBD segments, limiting the utility of segment number and length for genealogical inference (Smith et al., 2021).

Here, we present a new method, which we call specific parent-offspring relationship estimation (SPORE), that is robust to high levels of inbreeding and errors in genotyping. It is designed to use genome-wide genotypes from a variety of sources, such as those derived from high-coverage sequencing, imputation, or genotyping arrays. Instead of the commonly used IBD1 and IBD2, our method relies on related allele-sharing metrics that have large expected differences between parent-offspring and other types of relationships, even in highly inbred individuals. We show that SPORE is more sensitive and

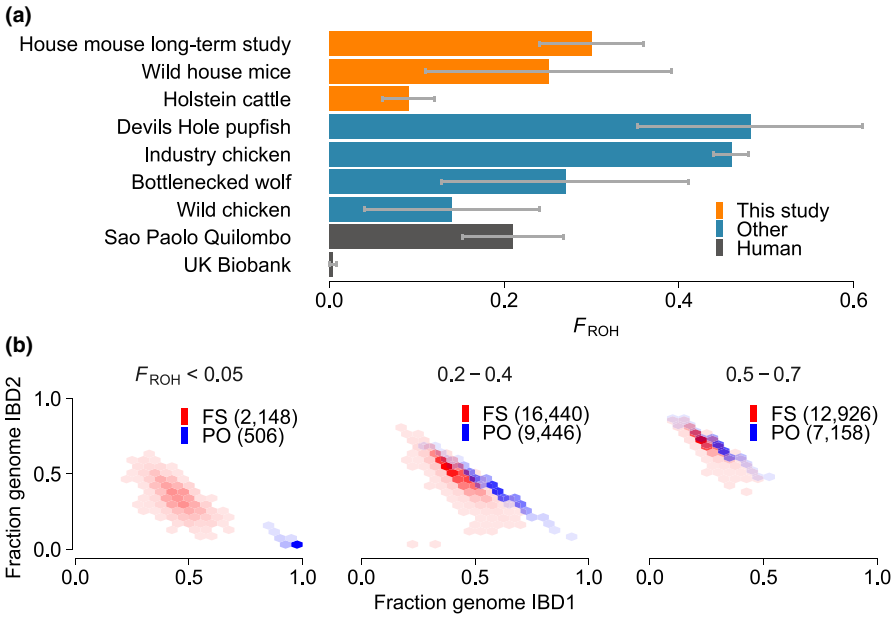

FIGURE 1 Extent of inbreeding in multiple populations and its influence on metrics used to infer relationships. (a) Overview of levels of inbreeding, measured as F_{ROH} (see Methods S1), in the populations studied here (orange) and in other nonhuman (blue) and human (grey) populations (Johnson et al., 2018; Kardos et al., 2018; Lemes et al., 2018; Talebi et al., 2020; Tian et al., 2021). Bars indicate average, error bars indicate the standard deviation, based on one population of the given group, with the exception of industry chicken (four groups). (b) Heatmaps of the fractions of genomes that are IBD1 and IBD2 in parent-offspring (PO, blue) and full sibling (FS, red) relationships. Each plot corresponds to different F_{ROH} levels. Data is based on the simulated pedigrees with genotyping errors (see Methods S1), with 17,110 PO and 31,514 FS relationships.

accurate than other commonly used methods that are not designed for inbred populations and that SPORE is more tolerant to genotyping errors, making it a robust tool for detecting relatives and building pedigrees using genetic data.

2 | MATERIALS AND METHODS

We developed SPORE as a robust method for parent-offspring inference and pedigree building in inbred populations and to be tolerant to genotyping errors. The main differences with existing methods are (1) the focus on variables that are expected to be zero in parent-offspring relationships but higher in other relationships (with the exception of monozygotic twins) and (2) the automatic adjustment of thresholds below which a relationship is classified as parent-offspring. This contrasts with other methods' use of fixed thresholds derived from theory assuming outbred populations (e.g., Manichaikul et al., 2010; Qiao et al., 2021), which can be violated by inbreeding and genotyping errors. Our approach also contrasts with COANCESTRY, which simulates parent-offspring relationships using the genetic information of the individuals in the data set, an approach that does not model genotyping errors nor resemble the distribution of inbreeding levels of the sample (Wang, 2011).

The essential elements of SPORE are described here, beginning with the variables that we use for finding parent-offspring relationships, followed by a description of the algorithm. Additional details are described in the Supporting Information.

SPORE takes a VCF file as input and its primary output is a pedigree file with the first columns in PED-compatible format. The IBD software we employ (TRUFFLE, Dimitromanolakis et al., 2019) requires ≥ 10,000 variants mapped to a genome reference. However, SPORE is modular and users can provide it with IBD values estimated through alternative methods.

2.1 | Variables used to infer parent-offspring relationships

To infer parent-offspring (PO) relationships from genome-wide genotypes of a set of individuals, we use three variables (see Figure 2), which we estimate for each pair of individuals: the proportion of the genome that is not IBD ("IBD0"), the variation (interquartile range; IQR) in that proportion among chromosomes ("IBD0 IQR"), and the proportion of loci that carry opposite homozygous genotypes ("homozygous mismatches, HM"). In outbred families, these variables have an expected value of zero in parent-offspring relationships (and in monozygotic twins), but higher than zero in full siblings and other types of relationships. IBD0 is expected to be invariably zero in all autosomes of parent-offspring pairs, yielding a total IBD0 of zero. In contrast, in full siblings on average $\frac{1}{4}$ of their autosomes are IBD2, $\frac{1}{2}$ IBD1, and $\frac{1}{4}$ IBD0. However, because of recombination and the independent assortment of chromosomes, this extent of sharing varies between chromosomes, and hence IBD0 IQR will be greater than zero. Notably, interchromosomal variation in IBD0 has not previously been leveraged as a metric to infer genealogical relatives (Huff et al., 2011).

Whereas IBD0, IBD0 IQR, and HM have expected values of zero in outbred parent-offspring relationships and higher than zero in other types of relationships, these metrics also move towards zero in full-siblings and other relationships under increased inbreeding (Figure 3a). Furthermore, genotyping errors can result in nonzero values for these metrics in parent-offspring relationships even in outbred individuals. Hence, by contrast to other methods developed for outbred individuals (e.g., Manichaikul et al., 2010; Qiao et al., 2021), we do not rely on hard-set thresholds below which we consider relationships to be parent-offspring. Instead, we use the information in the data set and a researcher-derived input variable—the assumed average parent-offspring relationships per individual (APO)—to automatically define those thresholds. Furthermore, we make use of a powerful metric that is also expected to be zero in parent-offspring relationships: the fraction of loci that carry genotypes that could not have been transmitted from a given set of two parents to their offspring in the absence of germline or embryonic de novo mutations (e.g., an A/A offspring with A/G and G/G putative parents), called "Mendelian trio errors," to determine correct father–mother-offspring trios. This metric performs well even under high inbreeding (Figure 3b), but is limited to cases where full trios are present in the sample. Similarly to IBD0, IBD0 IQR, and homozygous mismatches, we also automatically define thresholds below which Mendelian trio errors are considered low enough to be compatible with a true trio. Homozygous mismatches and Mendelian trio errors, like other identity-by-state metrics used in alternative genealogical relationship inference methods, are sensitive to SNP ascertainment biases of some genotyping approaches (Waples et al., 2019). However, the additional use of variables derived from called IBD0 segments in SPORE ameliorates that problem.

We have written SPORE to be able to incorporate additional user-provided variables with an expected value of zero in parent-offspring relationships but higher than zero in other types of relationships. It is also possible to deactivate the use of any or all of the three variables described above. One could, for example, use additional computational resources to calculate R0 (Waples et al., 2019) with NGSRELATEV2 (Hanghøj et al., 2019) and incorporate this metric (see Supporting Information "Evaluation of R0 as an alternative parent-offspring detection metric") into SPORE.

2.2 | Algorithm

SPORE is divided into three phases: (1) Detection of putative parent-offspring (PO) relationships (Figure 4, steps 1–3), (2) inference of father–mother-offspring trios within the putative PO relationships (Figure 4, steps 4–7), and (3) inference of PO relationships in the remaining putative PO relationships that could not be assigned to

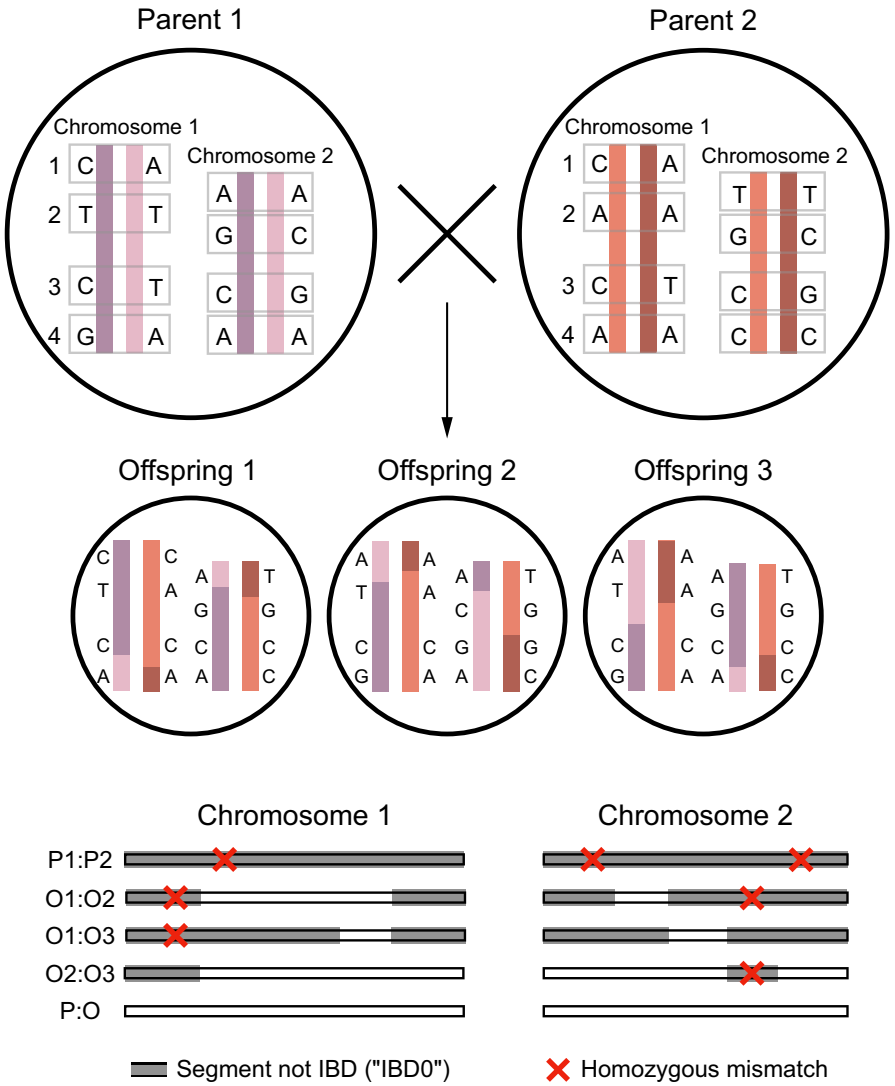

FIGURE 2 Schematic of metrics used by SPORE to infer parent-offspring relationships. Example of two unrelated parents who produce three offspring. Coloured chromosomes denote four different haplotypes (shades of purple and shades of orange), which are recombined and transmitted to three offspring (O1 to O3). Alleles are shown at four loci for each chromatid of two chromosome pairs. Bottom shows a comparison of the fraction of the genome that is not identical by descent ("IBD0," grey), the variation in IBD0 between chromosomes, and homozygous mismatches (red crosses) across pairs of individuals. Note that full-sibling (O:O) relationships differ in all three of these values from parent-offspring relationships (P:O), which are consistently 0: 0% of the genome IBD0, 0 variation in IBD0 ("IBD0 IQR"), 0 homozygous mismatches.

trios (Figure 4, steps 8–10). Details on SPORE's speed can be found in Figures S1 and S2.

2.2.1 | First phase

SPORE runs TRUFFLE 1.38 (Dimitromanolakis et al., 2019) to infer IBD0 between all possible combinations of individuals. Next, SPORE scans a random subset of the genome (Methods S1 and Figures S3 and S4) to detect homozygous mismatches (opposite genotypes) between individuals, which are divided by the number of loci at which both individuals are genotyped, resulting in a relative measure of homozygous mismatches ("HM").

Using IBD0, IBD0 IQR, and HM, SPORE aims to detect putative one-to-one PO relationships. To do so, SPORE uses the only necessary user input (apart from the genotypes themselves), the "assumed average PO relationships per individual" (APO) in the data set (see Performance of varying APO section). This input is used to automatically adjust the thresholds of IBD0, IBD0 IQR, and HM, below which a relationship is declared as putative PO (see Section 2.2.4 below).

2.2.2 | Second phase

SPORE aims to discover father–mother-offspring trios among the putative PO relationships. This phase can optionally be refined by providing birthdate and sex data. First, SPORE assembles putative father–mother-offspring trios out of the putative PO relationships for each focal individual that has at least two such relationships. SPORE then quantifies the number of loci that contain genotypes that could not have been transmitted from the putative parents to the putative offspring using the mendelian plugin of BCFTOOLS version 1.10.2 (Li, 2011). With those Mendelian trio errors, SPORE once again automatically estimates the thresholds below which a PO trio is considered to be true. If multiple trios pass the threshold(s), the one with fewest errors is assigned as true.

2.2.3 | Third phase

SPORE aims to infer PO relations in the remaining set of putative PO relationships that were so far not detected as a complete

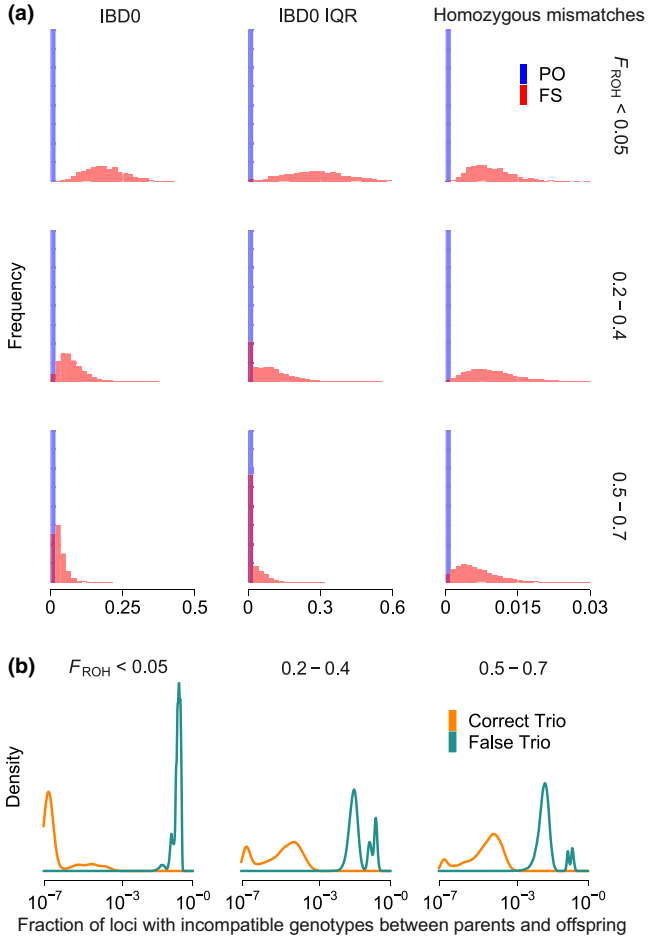

FIGURE 3 Overview of metrics used by SPORE to infer parent-offspring relationships. (a) Example normalized distributions of the variables that SPORE uses to distinguish parent-offspring (PO, blue) from full sibling (FS, red) relationships, across levels of inbreeding (F_{ROH}): The fraction of the genome that is not identical by descent ("IBD0"), the interchromosomal variation in IBD0 (measured as interquartile range) "IBD0 IQR," and the fraction of homozygous mismatches, inferred from 19 simulated autosomes, based on 82,911 PO and 120,291 FS relationships. (b) Example distributions of the fraction of incompatible genotypes (of up to ~2×10^6 loci) in 5909 true trios (orange) and in 111,976 putative, but with at least one incorrect, parents (green) across levels of inbreeding. Data in (a) and (b) are based on the simulated pedigrees with genotyping errors (see Methods S1). Note that the increase in errors of correct trios with increased inbreeding is an artefact of decreased genotype imputation accuracy using low-coverage sequencing for more inbred samples.

father–mother-offspring trio. Because these relationships can only be detected on a one-to-one rather than a more powerful trio basis, SPORE requires them to pass all three HM, IBD0, IBD0 IQR thresholds (see in Figure S5 how multiple thresholds decrease false positive PO detection). To be placed into the pedigree, these newly inferred putative PO relationships also need to fit into the existing trio-inferred pedigree. That means that one of the two individuals in a putative PO relationship needs to have unknown parents. If provided, the sex of the unknown parent and the birth order also need to be compatible. While we iterate through one remaining putative relationship at a time, we consider the alternatives for a given parent spot in the pedigree and choose the one with the lowest IBD0, IBD0 IQR, and HM values. Once this is done for all remaining putative PO relationships, SPORE returns the assembled pedigree. Putative PO relationships that pass all three thresholds but could not be placed into the assembled pedigree, for example because there was no known birthdate or because both directions (parent-offspring, offspring-parent) were possible for these two individuals, are output as "undirected PO relationships."

2.2.4 | Thresholding

Putative PO relationships

To establish thresholds below which relationships are considered as putative PO, SPORE calculates the number of individuals that have ≥ 1 PO relationship and > APO PO relationships at levels of IBD0, IBD0 IQR, and HM, separately. This is tested in up to 10,000 increments from the lowest value to the highest, with the first 9000 increments being exactly the 9000 lowest values, and the last 1000 increments being equally sized from the 9000th lowest value to the highest value. We do not exclude high values from this algorithm so SPORE remains flexible to various levels of genotyping errors at the cost of slightly longer runtime (see 3.3 Cattle population results). Fewer than 10,000 increments are used if fewer than 10,000 unique values exist of the variable that is looked at. The value at which the greatest distance between the fraction of individuals having at least one PO relationship and those that have more than the APO number of PO relationships is then set as the threshold of that variable below which relationships are classified as putative PO (Figure S6A). For example, if SPORE determined that 100% of individuals would have at least one relationship with IBD0 ≤ 10^{-5} and only 10% of individuals have more than the APO relationships at this threshold, then the distance would be 90 percentage points. If at any other threshold this distance is smaller than 90 percentage points, then 10^{-5} would be chosen as the IBD0 threshold for classifying putative PO relationships in this data set. At this stage, passing one threshold is enough to be classified as a putative PO relationship (see Figure S5 to see how multiple thresholds improve PO detection).

Putative father–mother-offspring trios

To determine the threshold below which Mendelian trio errors are low enough to consider the father–mother-offspring trio as true, Mendelian errors are used in two ways. In the first case, we aim to quantify how a trio's Mendelian errors compare to other trios that were evaluated for the same offspring. To that end, a trio's Mendelian errors are divided by the mean Mendelian trio errors of the other evaluated trios with the same putative offspring ("offspring-relative trio errors"). This could be useful if genotype quality varies between individuals, which could increase the errors in both the true and false trios. In the second case, we seek to analyse how low a trio's Mendelian errors are, compared to all evaluated trios (of all putative



FIGURE 4 Flowchart overview of the SPORE algorithm. input is shown as rectangles with square corners, while output is shown as rectangles with rounded corners. Details of each step can be found in the algorithm section.

offspring) across the data set. The resulting variable is simply the percentile of the relative error compared to all relative error counts ("data set-relative trio errors").

For both offspring- and data set-relative trio errors, SPORE then estimates the thresholds below which a trio is considered as true. Since only one set of parents can be true per offspring and only trios comprised of putative PO relationships are tested, SPORE makes use of two values to estimate the thresholds, (1) the fraction of putative offspring that would have exactly one set of parents, and (2) the fraction of putative offspring that would have more than one set of parents at each threshold. This is tested in up to 10,000 increments from the lowest value to the highest, with the first 9000 increments being exactly the 9000 lowest values, and the last 1000 increments being equally sized from the 9000th lowest value to the highest value. Fewer than 10,000 increments are used if fewer than 10,000 unique values exist of the variable being analysed. The threshold is then set at the trio-errors value at which the highest fraction of putative offspring are part of exactly one trio that falls below the threshold, but this fraction cannot be lower than the fraction of putative offspring that are part of more than one trio that passes the threshold (Figure S6B). For example, if 90% of putative offspring are part of exactly one trio that is below the fifth percentile of data set-relative trio errors and there is no threshold with a higher fraction of individuals with exactly two parents, then the fifth percentile is the threshold below which trios are considered to be true.

If only one trio was evaluated for a given putative offspring, only the data set-relative trio error threshold is used. When more than one trio pass a threshold for a putative offspring, that is, at least two sets of parents seem true, SPORE chooses the trio that has fewer errors. These cases are also output separately for the user to have the opportunity to examine more closely.

3 | RESULTS

To evaluate the performance of SPORE, we applied it to real and to simulated pedigrees of inbred populations. We then compared its performance to KING (Manichaikul et al., 2010), CREST (Qiao et al., 2021), and SEQUOIA (Huisman, 2017). Unlike SPORE, these methods are not explicitly designed for inbred populations and can also infer other non-PO relationships. KING provides very fast results in large data sets and is tolerant to population structure, with a major focus on human populations with no inbreeding. CREST's advantage is that it can determine more distant relationships, also under the assumption of no inbreeding. SEQUOIA requires only hundreds of SNPs to infer various types of relationships and is robust to some inbreeding. Like SPORE, CREST and SEQUOIA determine the direction (i.e., who is parent, who is offspring) of relationships, while KING does not. Table 1 presents an overview of these methods.

We used three data sets to compare the methods: (1) a subset of a population of wild house mice from a long-term study, (2) five simulated data sets based on the genotypes of the founders of the long-term house mice study, and (3) a published data set of a cattle pedigree and its genotypes (Alemu et al., 2021; Druet et al., 2020).

3.1 | Zurich house mouse population

The free-living wild house mice population has been intensively monitored since 2002, when it was founded using twelve wild-caught mice from two nearby source populations (König et al., 2015). The mice live in an old barn, which they can freely leave and re-enter. The barn is regularly searched for new litters and the mice are genotyped at 25 microsatellite loci when they are 13 days old and again as adults (> 17.5 g), and when they are found dead. The genotypes, along with information about the sex and dates of birth and death of individuals (which limits possible fathers to only those alive at the time pups were conceived and, in the case of mothers, when the pups were born) are used for pedigree construction. Only 79% of individuals can be confidently placed in the pedigree using this method (due to incomplete information from microsatellites and death dates), but to test SPORE we here make use only of individuals that have parents established with high certainty. We analyse a random subset of 151 mice plus their parents—204 mice in total as some of the 151 mice share parents or are parents themselves—of that population, chosen at random among the individuals that we have sequenced (see Methods S1). We imputed the genotypes of the mice using a custom analysis pipeline based on ANCESTRYHMM (Corbett-Detig & Nielsen, 2017; see Methods S1).

TABLE 1 Overview of the different algorithms compared with SPORE and some key differences between them with regards to parent-offspring inference

Measurement/feature	SPORE	CREST	KING	SEQUOIA
IBD	Yes, IBD0	Yes	Yes	No
IBD interchrom. Variation	Yes, IBD0	No	No	No
Kinship	No	No	Yes	Yes
Homozygous mismatches	Yes	No	No	Yes
Mendelian trio comparisons	Yes	No	No	Yes
Automatic thresholds	Yes	No	No	Via likelihood
Infers PO direction	Yes	Yes	No	Yes
Birthdates input	Optional	No	No	Optional
Sex input	Optional	No	No	Optional

3.1.1 | Performance

SPORE detected 295 of the 302 true parents (97.7%) correctly and assigned two parents incorrectly when birthdates were made available and 295 correct with four wrong assignments when birthdates were not used (Figure 5a). CREST did not find any parents. KING found 40% of parents, but also detected 88 false parents. SEQUOIA performed much better when it was given birthdate information. In that case, it found 32.5% of parents and only detected one wrong parent. Without birthdates, SEQUOIA found only 12.6% of parents while detecting seven wrong parents. Hence, SPORE found the highest number of correct parents, while making few mistakes.

3.2 | Simulated genotypes

We simulated five pedigrees (and the 19 autosomes of each individual) of randomly breeding mice of 50 overlapping generations based on the genotypes of the founders (across 2,161,810 variants) of the real house mouse population analysed in the previous section. There were on average 2369.2 (SD = 32.4) individuals per pedigree. Due to the overlapping of generations, offspring were at most at a pedigree depth of 35 equivalents of full generations, which is the unit we refer to in the analysis of specific generations. The pedigrees were simulated without any immigration and hence increasing levels of inbreeding to ~60% of their genome in runs of homozygosity (F_{ROH}) in the last generation (Figure S7). See Methods S1 for details.

3.2.1 | Performance

SPORE detected 93.2% ± 2.2 (mean ± SD) of the true parents (five simulations with $n = 4443 \pm 70.5$ true parents) correctly and detected on average 3.8 ± 1.8 wrong parents (Figure 5c). CREST inferred too many pairwise comparisons to be parent-offspring, so it did find 100% of parents, but only got the correct direction in 46% ± 0.8 of cases. It also found an average of 239,478 ± 67,089 wrong parents. KING found only 32.3% ± 5.6 of parents, and also detected 52 ± 9.6 false parents. SEQUOIA found 65.7% ± 3.8 of parents, and 144.2 ± 19.54 wrong parents. Overall, SPORE made the fewest mistakes, while still detecting more than 90% of true parents.

We also evaluated the impact of increased inbreeding, decreased genotyping quality, and incomplete population sampling through simulated pedigrees. To test performance on inbred individuals, we analysed pedigree generations 21–30 (Figure 6a). These individuals have a considerable level of inbreeding with an average of 49% of their genome in runs of homozygosity ($F_{ROH} = 0.49 \pm 0.07$; see Methods S1). SPORE's PO call accuracy in these generations was very similar to its accuracy over all generations (91% ± 3.4 vs. 93.2% detected true parents with 0.09% ± 0.09 vs. 0.09% of parent-offspring calls being wrong). CREST still detected too many false PO calls (104,496 ± 29,358). KING's performance was worse in these late generations than overall, detecting only 4.7% ± 4 versus 32.3% of true parents with 6.64% ± 5.8 versus 3.6% PO calls being wrong. SEQUOIA was the second most reliable algorithm after SPORE, with 48.8% ± 8.5 parents in generations 21–30 detected versus 65.7% parents in all

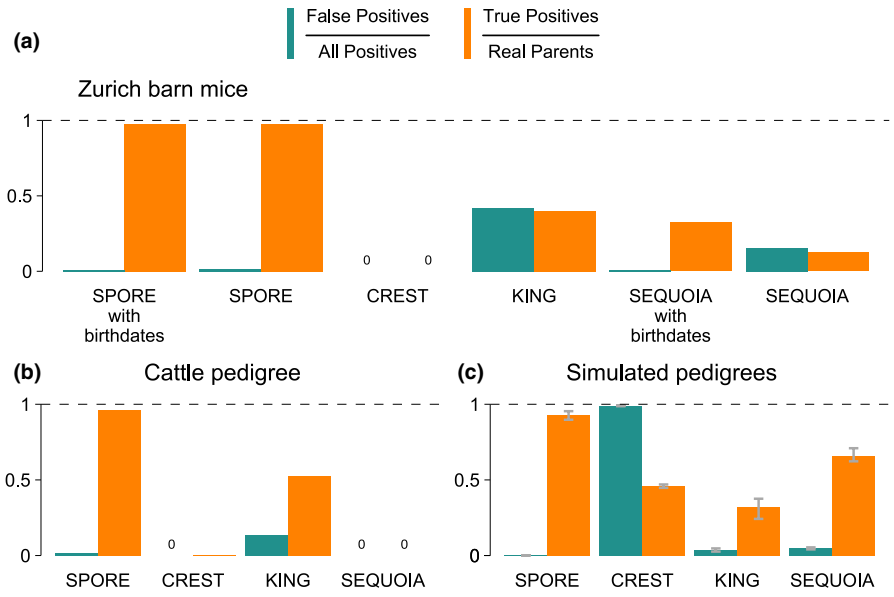

FIGURE 5 Performance of SPORE and three other pedigree inference methods. Two real populations are analysed: (a) house mice from the Zurich barn population (mean inbreeding level $F_{ROH} = 0.3 \pm 0.06$) and (b) a Holstein cattle pedigree (mean $F_{ROH} = 0.09 \pm 0.03$). (c) Performance on five simulated pedigrees (mean $F_{ROH} = 0.37 \pm 0.16$). Bar heights represent the mean of each simulation and error bars denote minimum and maximum. SPORE was run with APO = 6 in extensive-sampling mode (see Methods S1). SPORE, CREST, and SEQUOIA calls are only evaluated as true if the inferred direction of the call (who is parent, who is offspring) is correct or if the direction was not inferred.

generations and 10.1%±1.7 versus 4.7% false PO calls. Together, the results indicate that SPORE is superior at inferring PO correctly compared to other methods in controlled simulations with low and also with increasing levels of inbreeding.

With decreased genotyping quality (Figure 6b), but an otherwise unchanged data set, SPORE still found 84.5%±1.2 of true parents with 4.9%±2.1 false PO calls. In contrast, CREST now only found 0.02%±0.05 of the true parents with 92.5%±11.2 false PO calls. Similarly, KING was reduced to finding 0.3%±0.4 of true parents and 6.4%±8.8 false PO calls. SEQUOIA found 11.6%±1.3 of true parents with 81.8%±1.8 false PO calls. Thus, SPORE is considerably more robust to genotyping errors than other methods.

To disentangle the effects of inbreeding and genotyping error on the inference success of the different methods, we also inferred relationships in the first ten generations (which are the least inbred at F_{ROH} = 0.17±0.1) with and without decreased genotyping quality (Figure 6c,d). SPORE found 97.1%±0.7 of true parents without decreased genotyping quality and 86.3%±2.4 with decreased genotyping quality, whereas 0.1%±0.2 and 0%±0 of PO calls were wrong, respectively. In contrast, CREST's results changed from detecting too many PO relationships (78.6%±2.5 false PO calls) without decreased genotyping quality to detecting almost no true parents (0.03%±0.07 of true parents) with decreased genotyping quality. On the other hand, KING detected 80.7%±8.4 and 0.5%±1 of true parents with 2.7%±1.3 and 24.4%±43.3 wrong PO calls, respectively, showing a marked difference in performance between data sets with varying genotyping quality. Similarly, SEQUOIA detected 86.6%±4.7 of true parents with 0.9%±0.2 false PO calls without decreased genotyping quality, but only 21.7%±1.8 of true parents with 60.1%±4.2 false PO calls with decreased genotyping quality. These results are qualitatively similar to results based on individuals with minimal inbreeding ($F_{ROH} \leq 0.05$; Figure S8). In summary, SPORE is more robust to genotyping errors even under minimized inbreeding.

Sampling a random 75% of individuals in the data set (n = 2504±83.1 true parents) and hence no longer including the entire simulated pedigree, also affected the quality of inference (Figure 6e). SPORE, which works best the more full trios are in a data set, still found 97.9%±0.5 of true parents, but now also 1.1%±0.7 of PO calls were incorrect. CREST again detected too many false parents (136,718±36,558). KING found 33%±9 of true parents, but also 4.4%±1.4 false PO calls. SEQUOIA found 30.4%±4 of true parents and also 14.2%±3.5 false PO calls. When sampling is even less complete (50%), SPORE still detected 93.1%±3.1 of PO relationships but false positives also increased to 21.9%±8.5 (Figure S9). An alternative "intermediate sampling mode" allows SPORE to detect 86.6%±4.8 of PO relationships with only 0.2%±0.1 false PO calls in this scenario (see Figures S10–S12 for performance under varying sampling densities). CREST called essentially each pair as PO, while KING and SEQUOIA have a smaller increase in false positive calls than SPORE. Combining decreased genotyping quality with a 75% subset (Figure 6f) further decreased inference quality. Nevertheless, SPORE still found 81.4%±1.8 of true parents with 10.3%±3.6 false PO calls. By comparison, CREST only found 0.04%±0.06 of true parents with 70%±41.5 false PO calls. KING detected only 0.3%±0.34 of true parents, with 11.6%±11.2 false PO calls. SEQUOIA detected 13.1%±1.6 of true parents and 83.7%±2 wrong PO calls. Altogether, the results indicate that SPORE is a more accurate method when analysing inbred, incompletely sampled pedigrees even when there are abundant genotyping errors.

Performance with varying APO

To evaluate the impact of the user input "assumed parent-offspring relations per individual" (APO), we set it to a range from the lowest possible (1) to quintuple (30) of the value we used otherwise (6). Overall, SPORE detected more true parents at higher APO, but also slightly increased the false positives (Figure 7a). In the simulated pedigrees, the entire range of APO values delivered similar results, with at least 78±5.5% detected true parents and at most 2.6±2.1% false PO calls. However, APO had a greater impact on the quality of the inference in the house mouse study population: intermediate APO values led to similar results, while extreme values led to many missed or falsely inferred parents (Figure 7b). In sum, SPORE is not particularly sensitive to a range of APO but will have diminished accuracy at extreme values.

SPORE's APO-based thresholding is also robust to reproductive skew between the sexes, with APO best set at the assumed average PO relations overall, independent of sex, or a low value in general (see Supporting Information section APO under reproductive skew).

3.3 | Cattle population

To test SPORE's performance in an independent real pedigree, we applied it to a published set of 150 cattle offspring and their 100 unique parents and some grandparents, resulting in an incomplete pedigree with 320 known PO relationships (Alemu et al., 2021; Druet et al., 2020). We found high IBD0 values between parents and offspring (mean = 0.5; SD = 0.17; expected mean and SD = 0), suggesting a high genotyping error rate that would allow us to test SPORE's performance under these conditions.

3.3.1 | Performance

SPORE detected 96.3% of the true parents (n = 320) correctly and detected six wrong parents (Figure 5b). CREST found 0.3% of parents, with 0 false parents. KING found 52.8% of parents, but also detected 27 false parents. SEQUOIA did not find any parents. Thus, SPORE is superior at detecting parents in this independent real pedigree with high genotyping error rate.

4 | DISCUSSION

Our results based on data from a real house mouse population, a real cattle pedigree, and simulated pedigrees, show that SPORE

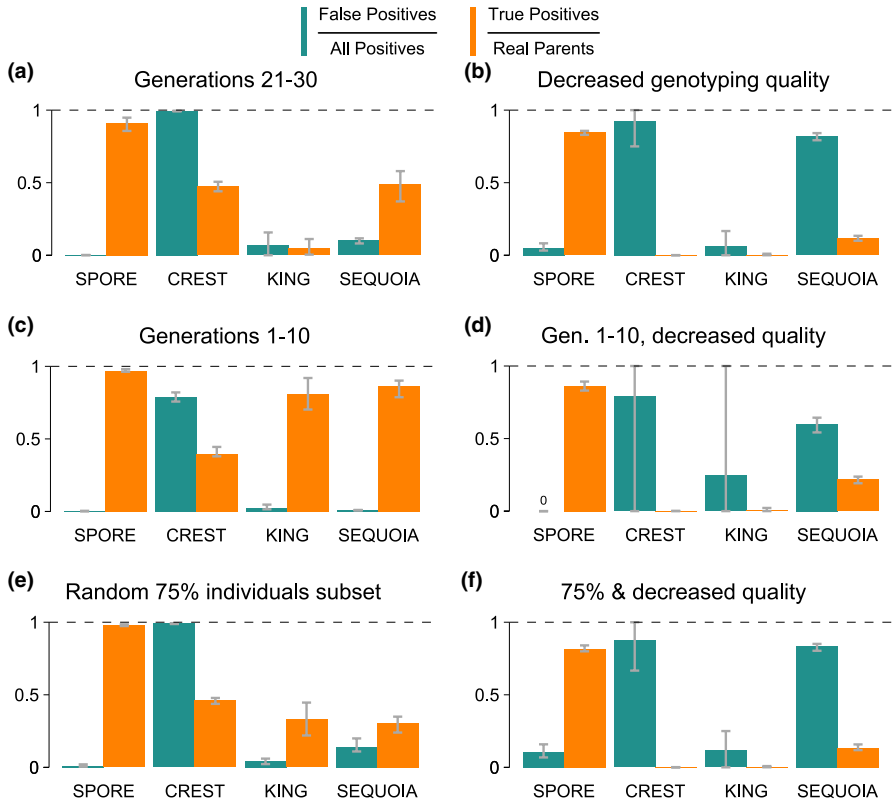

FIGURE 6 Performance of SPORE and three other algorithms under increased inbreeding, high genotyping errors and reduced pedigree sampling. Bar heights represent the mean of each simulation and error bars denote minimum and maximum. (a) Same simulated pedigrees as in Figure 5c but only parent-offspring calls of approximately the last 10 generations are analysed (mean $F_{ROH} = 0.49 \pm 0.07$). (b) Same pedigrees as in Figure 5c, but genotype imputation was performed with 10% simulated errors and 0.01× simulated coverage (rather than 2% and 0.03×). (c) the first 10 generations are analysed (mean $F_{ROH} = 0.17 \pm 0.1$). (d) a combination of (b) and (c), (e) same pedigrees and genotypes as Figure 5c, but only 75% of individuals were "sampled" (included in the analysis). (f) A combination of b and e. SPORE was run with APO = 6 in extensive-sampling mode. SPORE, CREST, and SEQUOIA calls are only evaluated as true if the inferred direction of the call (who is parent, who is offspring) is correct. SEQUOIA false positive calls only decrease by 5.7% (b) 11.4% (d), and 5.6% (f) when direction is ignored.

infers parent-offspring relationships in data sets with inbreeding or with high genotyping error, or both, with considerably more accuracy than other methods. The high performance of SPORE is based on (1) its reliance on an integrated combination of robust metrics, such as identity-by-descent of chromosomal segments, incompatible genotypes within trios and within parent-offspring pairs; (2) on the focus on parameters that have fixed expected values in parent-offspring relationships regardless of inbreeding levels; and (3) on the automatic adjustment of thresholds to be flexible to genotyping errors, which is a key difference to many other approaches.

Another advantage of SPORE is that it has few requirements. It does not need a genetic map and can optionally leverage sex information and birthdates. The minimal input—a VCF and a simple config file—also facilitates use by less bioinformatics-experienced investigators. More advanced use is also supported by the ability to incorporate additional metrics with fixed expected values in parent-offspring pairs, but not other relationships.

A caveat of SPORE is that it works on the assumption that there are parent-offspring trios in the analysed data set. When ~15% or more of individuals are part of a trio, false positive calls achieve an asymptotic low value < 1% (Figure S12). This assumption, however, is likely met in many of SPORE's intended uses. Furthermore, SPORE is designed to identify only parent-offspring relationships. However, parent-offspring inferences can then be used to identify siblings (individuals sharing parents) and to assemble multigenerational pedigrees. Pedigrees could be further completed using other methods such as SEQUOIA's full-sibling inference, although we have not evaluated its performance under inbreeding or imperfect genotypes.

We suggest that SPORE will be a useful tool for researchers monitoring experimental (e.g., Ferrari et al., 2022; Frentiu & Chenoweth, 2008; Luzynski et al., 2021; Stockley et al., 2013), wild (e.g., Bonnet et al., 2019; Chen et al., 2019; Clutton-Brock & Pemberton, 2003; Weinman et al., 2015), agricultural (e.g. McClure et al., 2018; Tortereau et al., 2017), and even some human (e.g., Arciero et al., 2021), populations with incomplete mating and birth tracking, potentially imperfect genotypes, and at least some autozygosity resulting from population history or recent mating between close relatives.

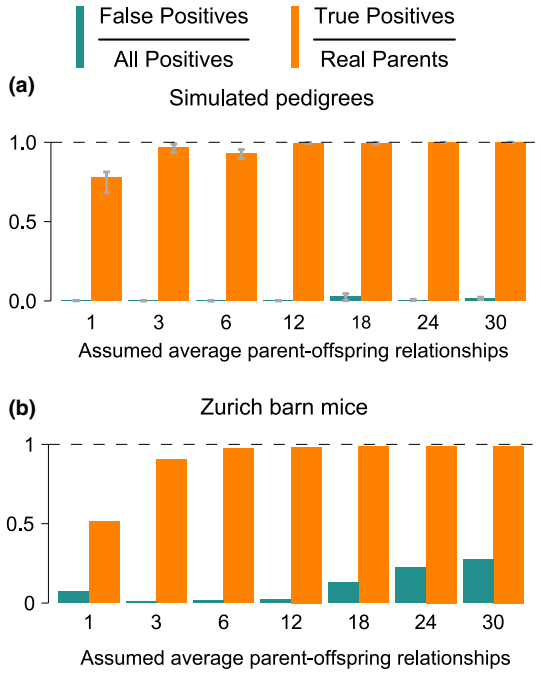

FIGURE 7 SPORE with different assumed average parent-offspring relationships per individual. SPORE was run in extensive-sampling mode. Six is the setting used for SPORE in all other plots. (a) Bar heights represent the mean of the five simulations and error bars denote minimum and maximum. (b) Results for the house mouse long-term study data set.

AUTHOR CONTRIBUTIONS

Conceptualization: Jan-Niklas Runge, Anna Lindholm, Barbara König and Andres Bendesky. Investigation: Jan-Niklas Runge and Anna Lindholm. Formal analysis: Jan-Niklas Runge. Methodology: Jan-Niklas Runge and Andres Bendesky. Resources: Anna K. Lindholm and Barbara König. Software: Jan-Niklas Runge. Visualization: Jan-Niklas Runge. Supervision: Andres Bendesky. Funding Acquisition: Jan-Niklas Runge, Anna K. Lindholm, Barbara König and Andres Bendesky. Writing: Jan-Niklas Runge, Anna K. Lindholm, Barbara König and Andres Bendesky.

ACKNOWLEDGEMENTS

Jari Garbely extracted DNA of house mouse samples and Kerel Francis prepared sequencing libraries. Sally Steinert, Bruce Boatman, and many others helped monitor and sample the mice in the barn of the long-term study. Itsik Pe'er provided comments on the manuscript. Open access funding provided by Universitat Zurich.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

FUNDING INFORMATION

Swiss National Science Foundation grants P1ZHP3_181303 & P2ZHP3_195249 (JNR), 31003A_176114 (BK), 31003A_120444 & 31003A_160328 (AL). Searle Scholarship, Klingenstein-Simons Fellowship, Sloan Foundation Fellowship, and National Institutes of Health award R35GM143051 (AB).

DATA AVAILABILITY STATEMENT

Sequence reads of the house mice from the long-term study have been made available in NCBI Sequence Read Archive under BioProject PRJNA782421. Founder genotypes used for imputation and simulation are available in Zenodo under doi 10.5281/zenodo.6465088. The cattle pedigree genotypes are available in Dryad under doi 10.5061/dryad.vx0k6djq8 (Druet et al., 2020). The SPORE version used here is archived in Zenodo under doi 10.5281/zenodo.6465798. SPORE is available at https://github.com/jnrunge/spore.

ORCID

Jan-Niklas Runge https://orcid.org/0000-0002-0450-9897
Barbara König https://orcid.org/0000-0001-7696-4736
Anna K. Lindholm https://orcid.org/0000-0001-8460-9769
Andres Bendesky https://orcid.org/0000-0002-4220-3653

REFERENCES

Alemu, S. W., Kadri, N. K., Harland, C., Faux, P., Charlier, C., Caballero, A., & Druet, T. (2021). An evaluation of inbreeding measures using a whole-genome sequenced cattle pedigree. *Heredity*, *126*(3), 410–423. https://doi.org/10.1038/s41437-020-00383-9

Arciero, E., Dogra, S. A., Malawsky, D. S., Mezzavilla, M., Tsismentzoglou, T., Huang, Q. Q., Hunt, K. A., Mason, D., Sharif, S. M., van Heel, D. A., Sheridan, E., Wright, J., Small, N., Carmi, S., Iles, M. M., & Martin, H. C. (2021). Fine-scale population structure and demographic history of British Pakistanis. *Nature Communications*, *12*(1), 7189. https://doi.org/10.1038/s41467-021-27394-2

Bonnet, T., Morrissey, M. B., Morris, A., Morris, S., Clutton-Brock, T. H., Pemberton, J. M., & Kruuk, L. E. B. (2019). The role of selection and evolution in changing parturition date in a red deer population. *PLoS Biology*, *17*(11), e3000493. https://doi.org/10.1371/journal.pbio.3000493

Bresadola, L., Link, V., Buerkle, C. A., Lexer, C., & Wegmann, D. (2020). Estimating and accounting for genotyping errors in RAD-seq experiments. *Molecular Ecology Resources*, *20*(4), 856–870. https://doi.org/10.1111/1755-0998.13153

Broad Institute. (2019). PICARD *toolkit*. Broad Institute. http://broadinstitute.github.io/picard

Browning, B. L., & Browning, S. R. (2013). Detecting identity by descent and estimating genotype error rates in sequence data. *American Journal of Human Genetics*, *93*(5), 840–851. https://doi.org/10.1016/j.ajhg.2013.09.014

Ceballos, F. C., Joshi, P. K., Clark, D. W., Ramsay, M., & Wilson, J. F. (2018). Runs of homozygosity: Windows into population history and trait architecture. *Nature Reviews. Genetics*, *19*(4), 220–234. https://doi.org/10.1038/nrg.2017.109

Chen, N., Juric, I., Cosgrove, E. J., Bowman, R., Fitzpatrick, J. W., Schoech, S. J., Clark, A. G., & Coop, G. (2019). Allele frequency dynamics in a pedigreed natural population. *Proceedings of the National Academy of Sciences*, *116*(6), 2158–2164. https://doi.org/10.1073/pnas.1813852116

Clutton-Brock, T. H., & Pemberton, J. M. (Eds.). (2003). *Soay sheep: Dynamics and selection in an Island population*. Cambridge University Press. https://doi.org/10.1017/CBO9780511550669

Corbett-Detig, R., & Nielsen, R. (2017). A hidden Markov model approach for simultaneously estimating local ancestry and admixture time using next generation sequence data in samples of arbitrary ploidy. *PLoS Genetics*, *13*(1), 1–40. https://doi.org/10.1371/journal.pgen.1006529

Dimitromanolakis, A., Paterson, A. D., & Sun, L. (2019). Fast and accurate shared segment detection and relatedness

estimation in un-phased genetic data via TRUFFLE. *American Journal of Human Genetics*, *105*(1), 78–88. https://doi.org/10.1016/j.ajhg.2019.05.007

Druet, T., Alemu, S. W., Kadri, N. K., Faux, P., Harland, C., Charlier, C., & Caballero, A. (2020). An evaluation of inbreeding measures using a whole genome sequenced cattle pedigree. *Dryad*. https://doi.org/10.5061/dryad.vx0k6djq8

Farquharson, K. A., Hogg, C. J., & Grueber, C. E. (2017). Pedigree analysis reveals a generational decline in reproductive success of captive Tasmanian devil (*Sarcophilus harrisii*): Implications for captive management of threatened species. *The Journal of Heredity*, *108*(5), 488–495. https://doi.org/10.1093/jhered/esx030

Ferrari, M., Lindholm, A. K., Ozgul, A., Oli, M. K., & König, B. (2022). Cooperation by necessity: Condition- and density-dependent reproductive tactics of female house mice. *Communications Biology*, *5*(1), 1–10. https://doi.org/10.1038/s42003-022-03267-2

Frentiu, F. D., & Chenoweth, S. F. (2008). Polyandry and paternity skew in natural and experimental populations of *Drosophila serrata*. *Molecular Ecology*, *17*(6), 1589–1596. https://doi.org/10.1111/j.1365-294X.2008.03693.x

Grant, P. R., & Grant, B. R. (2008). Pedigrees, assortative mating and speciation in Darwin's finches. *Proceedings of the Royal Society B: Biological Sciences*, *275*(1635), 661–668. https://doi.org/10.1098/rspb.2007.0898

Hanghøj, K., Moltke, I., Andersen, P. A., Manica, A., & Korneliussen, T. S. (2019). Fast and accurate relatedness estimation from high-throughput sequencing data in the presence of inbreeding. *GigaScience*, *8*(5), giz034. https://doi.org/10.1093/gigascience/giz034

Howe, L. J., Nivard, M. G., Morris, T. T., Hansen, A. F., Rasheed, H., Cho, Y., Chittoor, G., Lind, P. A., Palviainen, T., van der Zee, M. D., Cheesman, R., Mangino, M., Wang, Y., Li, S., Klaric, L., Ratliff, S. M., Bielak, L. F., Nygaard, M., Reynolds, C. A., … Davies, N. M. (2021). Within-sibship GWAS improve estimates of direct genetic effects. *bioRxiv*, 2021.03.05.433935. https://doi.org/10.1101/2021.03.05.433935

Huff, C. D., Witherspoon, D. J., Simonson, T. S., Xing, J., Watkins, W. S., Zhang, Y., Tuohy, T. M., Neklason, D. W., Burt, R. W., Guthery, S. L., Woodward, S. R., & Jorde, L. B. (2011). Maximum-likelihood estimation of recent shared ancestry (ERSA). *Genome Research*, *21*(5), 768–774. https://doi.org/10.1101/gr.115972.110

Huisman, J. (2017). Pedigree reconstruction from SNP data: Parentage assignment, sibship clustering and beyond. *Molecular Ecology Resources*, *17*(5), 1009–1024. https://doi.org/10.1111/1755-0998.12665

Johnson, E. C., Evans, L. M., & Keller, M. C. (2018). Relationships between estimated autozygosity and complex traits in the UKbiobank. *PLoS Genetics*, *14*(7), e1007556. https://doi.org/10.1371/journal.pgen.1007556

Kaplanis, J., Gordon, A., Shor, T., Weissbrod, O., Geiger, D., Wahl, M., Gershovits, M., Markus, B., Sheikh, M., Gymrek, M., Bhatia, G., MacArthur, D. G., Price, A. L., & Erlich, Y. (2018). Quantitative analysis of population-scale family trees with millions of relatives. *Science*, *360*(6385), 171–175. https://doi.org/10.1126/science.aam9309

Kardos, M., Åkesson, M., Fountain, T., Flagstad, Ø., Liberg, O., Olason, P., Sand, H., Wabakken, P., Wikenros, C., & Ellegren, H. (2018). Genomic consequences of intensive inbreeding in an isolated wolf population. *Nature Ecology and Evolution*, *2*(1), 124–131. https://doi.org/10.1038/s41559-017-0375-4

Keller, M. C., Visscher, P. M., & Goddard, M. E. (2011). Quantification of inbreeding due to distant ancestors and its detection using dense single nucleotide polymorphism data. *Genetics*, *189*(1), 237–249. https://doi.org/10.1534/genetics.111.130922

Koch, E. M., Schweizer, R. M., Schweizer, T. M., Stahler, D. R., Smith, D. W., Wayne, R. K., & Novembre, J. (2019). De novo mutation rate estimation in wolves of known pedigree. *Molecular Biology and Evolution*, *36*(11), 2536–2547. https://doi.org/10.1093/molbev/msz159

König, B., Lindholm, A. K., Lopes, P. C., Dobay, A., Steinert, S., & Buschmann, F. J.-U. (2015). A system for automatic recording of social behavior in a free-living wild house mouse population. *Animal Biotelemetry*, *3*(1), 39. https://doi.org/10.1186/s40317-015-0069-0

Lemes, R. B., Nunes, K., Carnavalli, J. E. P., Kimura, L., Mingroni-Netto, R. C., Meyer, D., & Otto, P. A. (2018). Inbreeding estimates in human populations: Applying new approaches to an admixed Brazilian isolate. *PLoS One*, *13*(4), e0196360. https://doi.org/10.1371/journal.pone.0196360

Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*, *27*(21), 2987–2993. https://doi.org/10.1093/bioinformatics/btr509

Li, H., Glusman, G., Hu, H., Shankaracharya, Caballero, J., Hubley, R., Witherspoon, D., Guthery, S. L., Mauldin, D. E., Jorde, L. B., Hood, L., Roach, J. C., & Huff, C. D. (2014). Relationship estimation from whole-genome sequence data. *PLoS Genetics*, *10*(1), e1004144. https://doi.org/10.1371/journal.pgen.1004144

Luzynski, K. C., Nicolakis, D., Marconi, M. A., Zala, S. M., Kwak, J., & Penn, D. J. (2021). Pheromones that correlate with reproductive success in competitive conditions. *Scientific Reports*, *11*(1), 21970. https://doi.org/10.1038/s41598-021-01507-9

Ma, L., O'Connell, J. R., VanRaden, P. M., Shen, B., Padhi, A., Sun, C., Bickhart, D. M., Cole, J. B., Null, D. J., Liu, G. E., Da, Y., & Wiggans, G. R. (2015). Cattle sex-specific recombination and genetic control from a large pedigree analysis. *PLoS Genetics*, *11*(11), e1005387. https://doi.org/10.1371/journal.pgen.1005387

Manichaikul, A., Mychaleckyj, J. C., Rich, S. S., Daly, K., Sale, M., & Chen, W.-M. (2010). Robust relationship inference in genome-wide association studies. *Bioinformatics*, *26*(22), 2867–2873. https://doi.org/10.1093/bioinformatics/btq559

McClure, M. C., McCarthy, J., Flynn, P., McClure, J. C., Dair, E., O'Connell, D. K., & Kearney, J. F. (2018). SNP data quality control in a national beef and dairy cattle system and highly accurate SNP based parentage verification and identification. *Frontiers in Genetics*, *9*, 84. https://doi.org/10.3389/fgene.2018.00084

Money, D., Gardner, K., Migicovsky, Z., Schwaninger, H., Zhong, G.-Y., & Myles, S. (2015). LinkImpute: Fast and accurate genotype imputation for nonmodel organisms. *G3: Genes, Genomes, Genetics*, *5*(11), 2383–2390. https://doi.org/10.1534/g3.115.021667

Ott, J., Wang, J., & Leal, S. M. (2015). Genetic linkage analysis in the age of whole-genome sequencing. *Nature Reviews. Genetics*, *16*(5), 275–284. https://doi.org/10.1038/nrg3908

Qiao, Y., Sannerud, J. G., Basu-Roy, S., Hayward, C., & Williams, A. L. (2021). Distinguishing pedigree relationships via multi-way identity by descent sharing and sex-specific genetic maps. *American Journal of Human Genetics*, *108*(1), 68–83. https://doi.org/10.1016/j.ajhg.2020.12.004

Smith, J., Qiao, Y., & Williams, A. L. (2021). Evaluating the utility of identity-by-descent segment numbers for relatedness inference via information theory and classification. *bioRxiv*, 2021.09.14.460357. https://doi.org/10.1101/2021.09.14.460357

Spielman, R. S., & Ewens, W. J. (1998). A sibship test for linkage in the presence of association: The sib transmission/disequilibrium test. *American Journal of Human Genetics*, *62*(2), 450–458. https://doi.org/10.1086/301714

Staples, J., Qiao, D., Cho, M. H., Silverman, E. K., Nickerson, D. A., & Below, J. E. (2014). PRIMUS: Rapid reconstruction of pedigrees from genome-wide estimates of identity by descent. *American Journal of Human Genetics*, *95*(5), 553–564. https://doi.org/10.1016/j.ajhg.2014.10.005

Stevens, E. L., Heckenberg, G., Roberson, E. D. O., Baugher, J. D., Downey, T. J., & Pevsner, J. (2011). Inference of relationships in population

data using identity-by-descent and identity-by-state. *PLoS Genetics*, *7*(9), e1002287. https://doi.org/10.1371/journal.pgen.1002287

Stockley, P., Ramm, S. A., Sherborne, A. L., Thom, M. D. F., Paterson, S., & Hurst, J. L. (2013). Baculum morphology predicts reproductive success of male house mice under sexual selection. *BMC Biology*, *11*(1), 66. https://doi.org/10.1186/1741-7007-11-66

Talebi, R., Szmatoła, T., Mészáros, G., & Qanbari, S. (2020). Runs of homozygosity in modern chicken revealed by sequence data. *G3: Genes, Genomes, Genetics*, *10*(12), 4615–4623. https://doi.org/10.1534/g3.120.401860

Tian, D., Turner, B. J., & Martin, C. H. (2021). Severe inbreeding and gene loss in the historical and extant population of the critically endangered Devils Hole pupfish. *bioRxiv*, 2021.08.13.456274. https://doi.org/10.1101/2021.08.13.456274

Tortereau, F., Moreno, C. R., Tosser-Klopp, G., Servin, B., & Raoul, J. (2017). Development of a SNP panel dedicated to parentage assignment in French sheep populations. *BMC Genetics*, *18*(1), 50. https://doi.org/10.1186/s12863-017-0518-2

Wang, J. (2004). Sibship reconstruction from genetic data with typing errors. *Genetics*, *166*(4), 1963–1979. https://doi.org/10.1093/genetics/166.4.1963

Wang, J. (2010). Effects of genotyping errors on parentage exclusion analysis. *Molecular Ecology*, *19*(22), 5061–5078. https://doi.org/10.1111/j.1365-294X.2010.04865.x

Wang, J. (2011). COANCESTRY: A program for simulating, estimating and analysing relatedness and inbreeding coefficients. *Molecular Ecology Resources*, *11*(1), 141–145. https://doi.org/10.1111/j.1755-0998.2010.02885.x

Waples, R. K., Albrechtsen, A., & Moltke, I. (2019). Allele frequency-free inference of close familial relationships from genotypes or low-depth sequencing data. *Molecular Ecology*, *28*(1), 35–48. https://doi.org/10.1111/mec.14954

Weinman, L. R., Solomon, J. W., & Rubenstein, D. R. (2015). A comparison of single nucleotide polymorphism and microsatellite markers for analysis of parentage and kinship in a cooperatively breeding bird. *Molecular Ecology Resources*, *15*(3), 502–511. https://doi.org/10.1111/1755-0998.12330

Wellmann, R. (2021). OPTISEL: Optimum contribution selection and population genetics. https://CRAN.R-project.org/package=optiSel

Wilson, A. J., Réale, D., Clements, M. N., Morrissey, M. M., Postma, E., Walling, C. A., Kruuk, L. E. B., & Nussey, D. H. (2010). An ecologist's guide to the animal model. *The Journal of Animal Ecology*, *79*(1), 13–26. https://doi.org/10.1111/j.1365-2656.2009.01639.x

Yengo, L., Wray, N. R., & Visscher, P. M. (2019). Extreme inbreeding in a European ancestry sample from the contemporary UK population. *Nature Communications*, *10*(1), 3719. https://doi.org/10.1038/s41467-019-11724-6

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Runge, J-N, König, B., Lindholm, A. K., & Bendesky, A. (2022). Parent-offspring inference in inbred populations. *Molecular Ecology Resources*, *00*, 1–13. https://doi.org/10.1111/1755-0998.13680